



#6

GENERAL INFORMATION

(iii) NUMBER OF SEQUENCES: 8

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2086 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2086

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGA TCC GGC CGT AGG AGA AGG AGT GTT CAG TGG TGC GCC GTA TCC CAA	48
Gly Ser Gly Arg Arg Arg Arg Ser Val Gln Trp Cys Ala Val Ser Gln	
1 5 10 15	
CCC GAG GCC ACA AAA TGC TTC CAA TGG CAA AGG AAT ATG AGA AAA GTG	96
Pro Glu Ala Thr Lys Cys Phe Gln Trp Gln Arg Asn Met Arg Lys Val	
20 25 30	
CGT GGC CCT CCT GTC AGC TGC ATA AAG AGA GAC TCC CCC ATC CAG TGT	144
Arg Gly Pro Pro Val Ser Cys Ile Lys Arg Asp Ser Pro Ile Gln Cys	
35 40 45	
ATC CAG GCC ATT GCG GAA AAC AGG GCC GAT GCT GTG ACC CTT GAT GGT	192
Ile Gln Ala Ile Ala Glu Asn Arg Ala Asp Ala Val Thr Leu Asp Gly	
50 55 60	
GGT TTC ATA TAC GAG GCA GGC CTG GCC CCC TAC AAA CTG CGA CCT GTA	240
Gly Phe Ile Tyr Glu Ala Gly Leu Ala Pro Tyr Lys Leu Arg Pro Val	
65 70 75 80	
GCG GCG GAA GTC TAC GGG ACC GAA AGA CAG CCA CGA ACT CAC TAT TAT	288
Ala Ala Glu Val Tyr Gly Thr Glu Arg Gln Pro Arg Thr His Tyr Tyr	
85 90 95	
GCC GTG GCT GTG GTG AAG AAG GGC GGC AGC TTT CAG CTG AAC GAA CTG	336
Ala Val Ala Val Val Lys Lys Gly Gly Ser Phe Gln Leu Asn Glu Leu	
100 105 110	
CAA GGT CTG AAG TCC TGC CAC ACA GGC CTT CGC AGG ACC GCT GGA TGG	384
Gln Gly Leu Lys Ser Cys His Thr Gly Leu Arg Arg Thr Ala Gly Trp	
115 120 125	

AAT GTC CCT ATA GGG ACA CTT CGT CCA TTC TTG AAT TGG ACG GGT CCA Asn Val Pro Ile Gly Thr Leu Arg Pro Phe Leu Asn Trp Thr Gly Pro 130 135 140	432
CCT GAG CCC ATT GAG GCA GCT GTG GCC AGG TTC TTC TCA GCC AGC TGT Pro Glu Pro Ile Glu Ala Ala Val Ala Arg Phe Phe Ser Ala Ser Cys 145 150 155 160	480
GTT CCC GGT GCA GAT AAA GGA CAG TTC CCC AAC CTG TGT CGC CTG TGT Val Pro Gly Ala Asp Lys Gly Gln Phe Pro Asn Leu Cys Arg Leu Cys 165 170 175	528
GCG GGG ACA GGG GAA AAC AAA TGT GCC TTC TCC TCC CAG GAA CCG TAC Ala Gly Thr Gly Glu Asn Lys Cys Ala Phe Ser Ser Gln Glu Pro Tyr 180 185 190	576
TTC AGC TAC TCT GGT GCC TTC AAG TGT CTG AGA GAC GGG GCT GGA GAC Phe Ser Tyr Ser Gly Ala Phe Lys Cys Leu Arg Asp Gly Ala Gly Asp 195 200 205	624
GTG GCT TTT ATC AGA GAG AGC ACA GTG TTT GAG GAC CTG TCA GAC GAG Val Ala Phe Ile Arg Glu Ser Thr Val Phe Glu Asp Leu Ser Asp Glu 210 215 220	672
GCT GAA AGG GAC GAG TAT GAG TTA CTC TGC CCA GAC AAC ACT CGG AAG Ala Glu Arg Asp Glu Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg Lys 225 230 235 240	720
CCA GTG GAC AAG TTC AAA GAC TGC CAT CTG GCC CGG GTC CCT TCT CAT Pro Val Asp Lys Phe Lys Asp Cys His Leu Ala Arg Val Pro Ser His 245 250 255	768
GCC GTT GTG GCA CGA AGT GTG AAT GGC AAG GAG GAT GCC ATC TGG AAT Ala Val Val Ala Arg Ser Val Asn Gly Lys Glu Asp Ala Ile Trp Asn 260 265 270	816
CTT CTC CGC CAG GCA CAG GAA AAG TTT GGA AAG GAC AAG TCA CCG AAA Leu Leu Arg Gln Ala Gln Glu Lys Phe Gly Lys Asp Lys Ser Pro Lys 275 280 285	864
TTC CAG CTC TTT GGC TCC CCT AGT GGG CAG AAA GAT CTG CTG TTC AAG Phe Gln Leu Phe Gly Ser Pro Ser Gly Gln Lys Asp Leu Leu Phe Lys 290 295 300	912
GAC TCT GCC ATT GGG TTT TCG AGG GTG CCC CCG AGG ATA GAT TCT GGG Asp Ser Ala Ile Gly Phe Ser Arg Val Pro Pro Arg Ile Asp Ser Gly 305 310 315 320	960
CTG TAC CTT GGC TCC GGC TAC TTC ACT GCC ATC CAG AAC TTG AGG AAA Leu Tyr Leu Gly Ser Gly Tyr Phe Thr Ala Ile Gln Asn Leu Arg Lys 325 330 335	1008
AGT GAG GAG GAA GTG GCT GCC CGG CGT GCG CGG GTC GTG TGG TGT GCG Ser Glu Glu Glu Val Ala Ala Arg Arg Ala Arg Val Val Trp Cys Ala 340 345 350	1056

GTG	GGC	GAG	CAG	GAG	CTG	CGC	AAG	TGT	AAC	CAG	TGG	AGT	GGC	TTG	AGC	1104
Val	Gly	Glu	Gln	Glu	Leu	Arg	Lys	Cys	Asn	Gln	Trp	Ser	Gly	Leu	Ser	
		355					360					365				
GAA	GGC	AGC	GTG	ACC	TGC	TCC	TCG	GCC	TCC	ACC	ACA	GAG	GAC	TGC	ATC	1152
Glu	Gly	Ser	Val	Thr	Cys	Ser	Ser	Ala	Ser	Thr	Thr	Glu	Asp	Cys	Ile	
	370					375					380					
GCC	CTG	GTG	CTG	AAA	GGA	GAA	GCT	GAT	GCC	ATG	AGT	TTG	GAT	GGA	GGA	1200
Ala	Leu	Val	Leu	Lys	Gly	Glu	Ala	Asp	Ala	Met	Ser	Leu	Asp	Gly	Gly	
385					390					395					400	
TAT	GTG	TAC	ACT	GCA	GGC	AAA	TGT	GGT	TTG	GTG	CCT	GTC	CTG	GCA	GAG	1248
Tyr	Val	Tyr	Thr	Ala	Gly	Lys	Cys	Gly	Leu	Val	Pro	Val	Leu	Ala	Glu	
			405					410						415		
AAC	TAC	AAA	TCC	CAA	CAA	AGC	AGT	GAC	CCT	GAT	CCT	AAC	TGT	GTG	GAT	1296
Asn	Tyr	Lys	Ser	Gln	Gln	Ser	Ser	Asp	Pro	Asp	Pro	Asn	Cys	Val	Asp	
		420						425					430			
AGA	CCT	GTG	GAA	GGA	TAT	CTT	GCT	GTG	GCG	GTG	GTT	AGG	AGA	TCA	GAC	1344
Arg	Pro	Val	Glu	Gly	Tyr	Leu	Ala	Val	Ala	Val	Val	Arg	Arg	Ser	Asp	
	435						440					445				
ACT	AGC	CTT	ACC	TGG	AAC	TCT	GTG	AAA	GGC	AAG	AAG	TCC	TGC	CAC	ACC	1392
Thr	Ser	Leu	Thr	Trp	Asn	Ser	Val	Lys	Gly	Lys	Lys	Ser	Cys	His	Thr	
	450					455					460					
GCC	GTG	GAC	AGG	ACT	GCA	GGC	TGG	AAT	ATC	CCC	ATG	GGC	CTG	CTC	TTC	1440
Ala	Val	Asp	Arg	Thr	Ala	Gly	Trp	Asn	Ile	Pro	Met	Gly	Leu	Leu	Phe	
465					470					475					480	
AAC	CAG	ACG	GGC	TCC	TGC	AAA	TTT	GAT	GAA	TAT	TTC	AGT	CAA	AGC	TGT	1488
Asn	Gln	Thr	Gly	Ser	Cys	Lys	Phe	Asp	Glu	Tyr	Phe	Ser	Gln	Ser	Cys	
			485					490						495		
GCC	CCT	GGG	TCT	GAC	CCG	AGA	TCT	AAT	CTC	TGT	GCT	CTG	TGT	ATT	GGC	1536
Ala	Pro	Gly	Ser	Asp	Pro	Arg	Ser	Asn	Leu	Cys	Ala	Leu	Cys	Ile	Gly	
		500						505					510			
GAC	GAG	CAG	GGT	GAG	AAT	AAG	TGC	GTG	CCC	AAC	AGC	AAC	GAG	AGA	TAC	1584
Asp	Glu	Gln	Gly	Glu	Asn	Lys	Cys	Val	Pro	Asn	Ser	Asn	Glu	Arg	Tyr	
		515					520					525				
TAC	GGC	TAC	ACT	GGG	GCT	TTC	CGG	TGC	CTG	GCT	GAG	AAT	GCT	GGA	GAC	1632
Tyr	Gly	Tyr	Thr	Gly	Ala	Phe	Arg	Cys	Leu	Ala	Glu	Asn	Ala	Gly	Asp	
	530					535					540					
GTT	GCA	TTT	GTG	AAA	GAT	GTC	ACT	GTC	TTG	CAG	AAC	ACT	GAT	GGA	AAT	1680
Val	Ala	Phe	Val	Lys	Asp	Val	Thr	Val	Leu	Gln	Asn	Thr	Asp	Gly	Asn	
545					550					555					560	
AAC	AAT	GAG	GCA	TGG	GCT	AAG	GAT	TTG	AAG	CTG	GCA	GAC	TTT	GCG	CTG	1728
Asn	Asn	Glu	Ala	Trp	Ala	Lys	Asp	Leu	Lys	Leu	Ala	Asp	Phe	Ala	Leu	
				565					570					575		
CTG	TGC	CTC	GAT	GGC	AAA	CGG	AAG	CCT	GTG	ACT	GAG	GCT	AGA	AGC	TGC	1776

Leu	Cys	Leu	Asp	Gly	Lys	Arg	Lys	Pro	Val	Thr	Glu	Ala	Arg	Ser	Cys	
			580					585					590			
CAT	CTT	GCC	ATG	GCC	CCG	AAT	CAT	GCC	GTG	GTG	TCT	CGG	ATG	GAT	AAG	1824
His	Leu	Ala	Met	Ala	Pro	Asn	His	Ala	Val	Val	Ser	Arg	Met	Asp	Lys	
		595					600					605				
GTG	GAA	CGC	CTG	AAA	CAG	GTG	TTG	CTC	CAC	CAA	CAG	GCT	AAA	TTT	GGG	1872
Val	Glu	Arg	Leu	Lys	Gln	Val	Leu	Leu	His	Gln	Gln	Ala	Lys	Phe	Gly	
	610					615					620					
AGA	AAT	GGA	TCT	GAC	TGC	CCG	GAC	AAG	TTT	TGC	TTA	TTC	CAG	TCT	GAA	1920
Arg	Asn	Gly	Ser	Asp	Cys	Pro	Asp	Lys	Phe	Cys	Leu	Phe	Gln	Ser	Glu	
625					630					635					640	
ACC	AAA	AAC	CTT	CTG	TTC	AAT	GAC	AAC	ACT	GAG	TGT	CTG	GCC	AGA	CTC	1968
Thr	Lys	Asn	Leu	Leu	Phe	Asn	Asp	Asn	Thr	Glu	Cys	Leu	Ala	Arg	Leu	
			645					650						655		
CAT	GGC	AAA	ACA	ACA	TAT	GAA	AAA	TAT	TTG	GGA	CCA	CAG	TAT	GTC	GCA	2016
His	Gly	Lys	Thr	Thr	Tyr	Glu	Lys	Tyr	Leu	Gly	Pro	Gln	Tyr	Val	Ala	
			660					665					670			
GGC	ATT	ACT	AAT	CTG	AAA	AAG	TGC	TCA	ACC	TCC	CCC	CTC	CTG	GAA	GCC	2064
Gly	Ile	Thr	Asn	Leu	Lys	Lys	Cys	Ser	Thr	Ser	Pro	Leu	Leu	Glu	Ala	
		675					680					685				
TGT	GAA	TTC	CTC	AGG	AAG	TAA	A									2086
Cys	Glu	Phe	Leu	Arg	Lys	*										
	690					695										

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 695 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gly	Ser	Gly	Arg	Arg	Arg	Arg	Ser	Val	Gln	Trp	Cys	Ala	Val	Ser	Gln	
1				5					10					15		
Pro	Glu	Ala	Thr	Lys	Cys	Phe	Gln	Trp	Gln	Arg	Asn	Met	Arg	Lys	Val	
			20					25					30			
Arg	Gly	Pro	Pro	Val	Ser	Cys	Ile	Lys	Arg	Asp	Ser	Pro	Ile	Gln	Cys	
		35					40					45				
Ile	Gln	Ala	Ile	Ala	Glu	Asn	Arg	Ala	Asp	Ala	Val	Thr	Leu	Asp	Gly	
	50					55					60					
Gly	Phe	Ile	Tyr	Glu	Ala	Gly	Leu	Ala	Pro	Tyr	Lys	Leu	Arg	Pro	Val	
65					70					75					80	

Ala Ala Glu Val Tyr Gly Thr Glu Arg Gln Pro Arg Thr His Tyr Tyr
 85 90 95
 Ala Val Ala Val Val Lys Lys Gly Gly Ser Phe Gln Leu Asn Glu Leu
 100 105 110
 Gln Gly Leu Lys Ser Cys His Thr Gly Leu Arg Arg Thr Ala Gly Trp
 115 120 125
 Asn Val Pro Ile Gly Thr Leu Arg Pro Phe Leu Asn Trp Thr Gly Pro
 130 135 140
 Pro Glu Pro Ile Glu Ala Ala Val Ala Arg Phe Phe Ser Ala Ser Cys
 145 150 155 160
 Val Pro Gly Ala Asp Lys Gly Gln Phe Pro Asn Leu Cys Arg Leu Cys
 165 170 175
 Ala Gly Thr Gly Glu Asn Lys Cys Ala Phe Ser Ser Gln Glu Pro Tyr
 180 185 190
 Phe Ser Tyr Ser Gly Ala Phe Lys Cys Leu Arg Asp Gly Ala Gly Asp
 195 200 205
 Val Ala Phe Ile Arg Glu Ser Thr Val Phe Glu Asp Leu Ser Asp Glu
 210 215 220
 Ala Glu Arg Asp Glu Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg Lys
 225 230 235 240
 Pro Val Asp Lys Phe Lys Asp Cys His Leu Ala Arg Val Pro Ser His
 245 250 255
 Ala Val Val Ala Arg Ser Val Asn Gly Lys Glu Asp Ala Ile Trp Asn
 260 265 270
 Leu Leu Arg Gln Ala Gln Glu Lys Phe Gly Lys Asp Lys Ser Pro Lys
 275 280 285
 Phe Gln Leu Phe Gly Ser Pro Ser Gly Gln Lys Asp Leu Leu Phe Lys
 290 295 300
 Asp Ser Ala Ile Gly Phe Ser Arg Val Pro Pro Arg Ile Asp Ser Gly
 305 310 315 320
 Leu Tyr Leu Gly Ser Gly Tyr Phe Thr Ala Ile Gln Asn Leu Arg Lys
 325 330 335
 Ser Glu Glu Glu Val Ala Ala Arg Arg Ala Arg Val Val Trp Cys Ala
 340 345 350
 Val Gly Glu Gln Glu Leu Arg Lys Cys Asn Gln Trp Ser Gly Leu Ser
 355 360 365
 Glu Gly Ser Val Thr Cys Ser Ser Ala Ser Thr Thr Glu Asp Cys Ile
 370 375 380

Ala	Leu	Val	Leu	Lys	Gly	Glu	Ala	Asp	Ala	Met	Ser	Leu	Asp	Gly	Gly	
385					390					395					400	
Tyr	Val	Tyr	Thr	Ala	Gly	Lys	Cys	Gly	Leu	Val	Pro	Val	Leu	Ala	Glu	
				405					410					415		
Asn	Tyr	Lys	Ser	Gln	Gln	Ser	Ser	Asp	Pro	Asp	Pro	Asn	Cys	Val	Asp	
			420					425					430			
Arg	Pro	Val	Glu	Gly	Tyr	Leu	Ala	Val	Ala	Val	Val	Arg	Arg	Ser	Asp	
		435					440						445			
Thr	Ser	Leu	Thr	Trp	Asn	Ser	Val	Lys	Gly	Lys	Lys	Ser	Cys	His	Thr	
	450					455						460				
Ala	Val	Asp	Arg	Thr	Ala	Gly	Trp	Asn	Ile	Pro	Met	Gly	Leu	Leu	Phe	
465					470					475					480	
Asn	Gln	Thr	Gly	Ser	Cys	Lys	Phe	Asp	Glu	Tyr	Phe	Ser	Gln	Ser	Cys	
				485					490					495		
Ala	Pro	Gly	Ser	Asp	Pro	Arg	Ser	Asn	Leu	Cys	Ala	Leu	Cys	Ile	Gly	
			500					505					510			
Asp	Glu	Gln	Gly	Glu	Asn	Lys	Cys	Val	Pro	Asn	Ser	Asn	Glu	Arg	Tyr	
		515					520						525			
Tyr	Gly	Tyr	Thr	Gly	Ala	Phe	Arg	Cys	Leu	Ala	Glu	Asn	Ala	Gly	Asp	
	530					535					540					
Val	Ala	Phe	Val	Lys	Asp	Val	Thr	Val	Leu	Gln	Asn	Thr	Asp	Gly	Asn	
545					550					555					560	
Asn	Asn	Glu	Ala	Trp	Ala	Lys	Asp	Leu	Lys	Leu	Ala	Asp	Phe	Ala	Leu	
				565					570					575		
Leu	Cys	Leu	Asp	Gly	Lys	Arg	Lys	Pro	Val	Thr	Glu	Ala	Arg	Ser	Cys	
			580					585					590			
His	Leu	Ala	Met	Ala	Pro	Asn	His	Ala	Val	Val	Ser	Arg	Met	Asp	Lys	
		595					600					605				
Val	Glu	Arg	Leu	Lys	Gln	Val	Leu	Leu	His	Gln	Gln	Ala	Lys	Phe	Gly	
	610					615						620				
Arg	Asn	Gly	Ser	Asp	Cys	Pro	Asp	Lys	Phe	Cys	Leu	Phe	Gln	Ser	Glu	
625					630					635					640	
Thr	Lys	Asn	Leu	Leu	Phe	Asn	Asp	Asn	Thr	Glu	Cys	Leu	Ala	Arg	Leu	
				645					650					655		
His	Gly	Lys	Thr	Thr	Tyr	Glu	Lys	Tyr	Leu	Gly	Pro	Gln	Tyr	Val	Ala	
			660					665					670			
Gly	Ile	Thr	Asn	Leu	Lys	Lys	Cys	Ser	Thr	Ser	Pro	Leu	Leu	Glu	Ala	
		675					680					685				

Cys Glu Phe Leu Arg Lys *
690 695

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCGGATCCG GCCGTAGGAG AAGGAGTGTT CAGTGG

36

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGATCTAGAT TACTTCCTGA GGAATCCACA GGC

33

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 42 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AAGCTTATGA AACTTGTCTT CCTCGTCCTG TTCTTCCTCG GG

42

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GATCCAGCCA GAGAGAGTCC GAGGGCCCCG AGGAA

35

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATTATTGAA ACGAGGAATT AGCTTATG

28

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TCGAGAAAAG ACTTTTCTCC GG

22